IN THE UNITED STATES DISTRICT COURT
FOR THE WESTERN DISTRICT OF PENNSYLVANIA

UNITED STATES OF AMERICA	)	
	)	
v.	)	Criminal No. 19-369
	)	
LAFON ELLIS	)	

DECLARATION OF DR. DAN E. KRANE

I, Dan E. Krane, Ph.D., declare I have personal knowledge of the following, and if called upon to do so, could and would testify competently to the matters contained herein.

I. Qualifications

1. I am the Interim Dean and Chief Administrative Officer of Wright State University's Lake Campus in Celina, Ohio. I am also a professor in the Department of Biological Sciences with a courtesy appointment in Computer Science at Wright State University in Dayton, Ohio. I have recently completed a one-year term as a fellow of the American Council on Education at the University of Notre Dame in Notre Dame, Indiana. I have a B.S. degree with a double major in Biology and Chemistry from John Carroll University (Cleveland, Ohio), and a Ph.D. from the Biochemistry program of the Cell and Molecular Biology Department of the Pennsylvania State University (State College, Pennsylvania). I have also done postdoctoral research using the tools of molecular biology to answer questions in the fields of population genetics and molecular evolution in the Genetics Department of the Washington University Medical School (St. Louis, Missouri) and in the Department of Organismic and Evolutionary Biology of Harvard University (Cambridge, Massachusetts). I have published more than 50 scholarly papers on a variety of topics including population genetic studies of the genetic diversity of human populations at DNA typing loci, of organisms exposed to environmental stressors, and the use of DNA typing in forensic science. I am also the lead author of a widely used undergraduate textbook, *Fundamental*

Exhibit 3

Concepts of Bioinformatics. I was a founding member of and two-time gubernatorial appointee to the Commonwealth of Virginia's Scientific Advisory Committee, a 12-member panel established by statute to provide oversight and guidance to the Virginia Department of Forensic Science (the crime laboratory for the Commonwealth of Virginia). I have testified in more than 110 criminal proceedings that have involved forensic DNA typing (in 23 different state and district courts, several courts martial, and in five different federal courts within the United States, a Coronial Inquest in the State of Victoria in Australia, in Belfast Crown Court in Northern Ireland, the Black Friar's Court and the Central Criminal Court of London, and in Oxford Crown Court in England). Included in those cases in which I have testified are instances where I have worked as an independent expert for a Coronial Inquest and in US federal court (W.D. Mich.).

II. Background

2. I have been asked by attorney Khasha Attaran to assist with the evaluation of results generated by TrueAllele® in the case of *United States v. Lafon Ellis* and specifically to comment on the utility of access to the TrueAllele® source code and associated software development materials. I have received a Cybergenetics TrueAllele® Report dated December 5, 2019 and materials underlying the TrueAllele® analysis generated during the course of testing Item 2A, swabbing of pistol. I have also received a Cybergenetics Supplemental Report dated October 23, 2020 describing results from additional analyses conducted by Cybergenetics on these same data but using five additional probabilistic genotyping (PG) software systems. I have reviewed results generated by TrueAllele® in many criminal cases. I am very familiar with the conventional interpretation of DNA test results (e.g. as done by human experts without the assistance of probabilistic genotyping

software like TrueAllele®). Access to source code can assist with determining the confidence with which PG results should be considered in two significant areas: comprehension of the system's behaviors and verification & validation (V&V) of those behaviors.

III. Results

3. The December 5, 2019 report states: "TrueAllele assumed that the evidence sample data (Item 2A) contained three, four, or five unknown contributors, and objectively inferred evidence genotypes solely from these data. Degraded DNA was considered." And, "A match between the pistol (Item 2A) and Lafon Ellis (Item 4) is: 21.4 trillion times more probable than a coincidental match to an unrelated African-American person, 9.42 quadrillion times more probable than a coincidental match to an unrelated Caucasian person, and 19.2 quadrillion times more probable than a coincidental match to an unrelated Hispanic person."

4. The October 23, 2020 report states, "The data were analyzed assuming four contributors. When the software was limited to under four contributors, drop-in was considered." Likelihood ratios (LRs) reported for comparisons of the reference profile for Lafon Ellis to Item 2A, swabbing of pistol, and using each of the five systems used are:

LRmix	1,440
Lab Retriever	110,000
likeLTD v5.5	36,700
likeLTD v6.3	721,000,000
EuroForMix	2,020,000,000,000,000

Table 1 – Derived from "DNA Match Tables" Table "1. Likelihood ratio" from Cybergenetics report dated October 23, 2020.

IV. Utility of source code in review

5. Mathematical or technical English descriptions of intended software behaviors are not

actual software behaviors. Translations from concept to source code often involve compromises due to a variety of factors, usually due to resource limitations. These factors can include time needed for a complex algorithm to calculate an optimal solution to a problem as opposed to a simpler algorithm that executes much more quickly to arrive at a "good enough" but sub-optimal solution. Another common factor are hardware limitations such as memory or CPU speed. A longstanding example of these compromises in the bioinformatics space is the Clustal family of DNA and protein sequence alignment software programs, which are provably optimal when aligning only two sequences but are usually sub-optimal when simultaneously aligning three or more sequences—a compromise made due to the computational complexity of the algorithms. For more than two decades, improvements have been iteratively made to further optimize the multiple sequence alignment algorithms used in the Clustal software, but it is still incapable of provably optimal results for multiple sequence alignments.

6. TrueAllele® uses a Markov chain Monte Carlo (MCMC) algorithm to sample variables when exploring possible combinations of contributor genotypes. MCMC algorithms are, by their sampling nature, a compromise away from optimal solutions due to computational complexity. While it might be economical to sample from all possible combinations of variables, an MCMC algorithm will not provide a provably optimal solution.

7. It is important to understand where other compromises exist within TrueAllele® due to time or hardware limitations. While the overarching MCMC algorithm used by TrueAllele® is an inherent compromise, a review of the actual implementation of the system, i.e. its source code, would provide insight into any other compromises made in subcomponents of its hierarchical model. An incomplete list of example subcomponents

include how TrueAllele® addresses the following concepts relevant to forensic DNA interpretation:

a) distinguishing between signal and noise,

b) recognizing and accounting for stutter artifacts,

c) peak smoothing,

d) peak height imbalance,

e) contributor proportions of total DNA,

f) inhibition and differential degradation,

g) drop-in, and

h) population genetics considerations such as relatedness or inbreeding.

8. While each of these concepts can and have been modeled and accounted for through different algorithms, each particular algorithm can be implemented variously on a spectrum between "optimally correct" and "fast." Comprehending the strengths and weaknesses of the actual implementation is important when considering how much confidence should be placed in that particular implementation.

9. It is worth noting that each of the five alternative PG programs used by Cybergenetics for its October 23, 2020 report are free and open-source software, accessible to anyone with internet access. Any hindrance to answering questions about the innerworkings of these programs is limited by comprehension of the biological, statistical, or programming concepts—not by lack of access to the source code. Given the range of twelve orders of magnitude in the LRs calculated by these programs, characterizing the sources of this variation is possible by not only hypothetically comparing the systems as published algorithms, but also by evaluating their specific, actual behaviors during the calculations.

10. On the subject of V&V, Christopher Steele and David Balding, co-developers of the likeLTD v6 software used in this case, have written:

> "Some progress can be made in evaluating the validity and performance of software. Courts need these kinds of evaluations to have confidence in the results of software-based forensic analyses. Open source software is highly desirable in the court environment because openness to scrutiny by any interested party is an invaluable source of bug reports and suggestions for improvement." ("Statistical evaluation of forensic DNA profile evidence," Annu. Rev. Stat. Its Appl., vol. 1, pp. 361–384, 2014.)

11. I agree with Christopher Steele and David Balding that openness allows for greater scrutiny.

12. In 2016, the International Society for Forensic Genetics (ISFG) published guidance on validating probabilistic genotyping software, including on "the minimum requirements for the validation (is it doing the right thing?) and verification (is it doing the thing right?) of a software program (V&V)…." (Coble, et al., "DNA Commission of the International Society for Forensic Genetics: Recommendations on the validation of software programs performing biostatistical calculations for forensic genetics applications," Forensic Sci. Int. Genet., vol. 25, pp. 191–197, 2016.) The authors cited software engineering standards as examples of what "can be simplified and extrapolated to forensic genetics" for probabilistic genotyping software development and validation.

13. In 2018, I co-authored a letter to the editor of the Journal of Forensic Sciences on the subject of V&V of probabilistic genotyping software, suggesting:

> "Software-based PG approaches are necessarily rooted in collaboration between experts in the areas of molecular biology, population genetics, statistics, forensic science, computer science, and software engineering. While it is important to consider the perspectives of all of these disciplines on the validation issue, we think that the perspectives of software engineers are particularly important." (N. Adams, et al., "Letter to the Editor-Appropriate Standards for Verification and Validation of Probabilistic Genotyping Systems," J. Forensic Sci., vol. 63, no. 1, pp. 339–340, 2018.)

14. Software engineering V&V principles and standards emphasize access to source code and other software development materials for review during various V&V tasks. As ISFG puts it, careful review of these materials allow for informed decisions as to whether the PG system is "doing the thing right" (verification) as well as "doing the right thing" (validation).

15. In 2016, the President's Council of Advisors on Science and Technology (PCAST) also recognized these concerns:

> "These probabilistic genotyping software programs clearly represent a major improvement over purely subjective interpretation. However, they still require careful scrutiny to determine (1) whether the methods are scientifically valid, including defining the limitations on their reliability (that is, the circumstances in which they may yield unreliable results) and (2) whether the software correctly implements the methods. This is particularly important because the programs employ different mathematical algorithms and can yield different results for the same mixture profile." (President's Council of Advisors on Science and Technology, "Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods," 2016. Available: https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensic_science_report_final.pdf)

16. PCAST also recognized that difficulties in mixture interpretation are complicated by many factors, including the number of contributors to a mixture: "DNA analysis of complex mixtures—defined as mixtures with more than two contributors—is inherently difficult and even more for small amounts of DNA."

17. Item 2A, swabbing of pistol, was interpreted by Cybergenetics as originating from three, four, or five contributors. Each of these interpretations would meet the PCAST criteria for a "complex mixture."

18. Since each of the PG software systems used in this case generated a different result for the comparison of the reference profile of Lafon Ellis to Item 2A, swabbing of pistol, none of

them can be used as an answer key or ground truth to confirm "whether the software correctly implements the methods." The most direct way to evaluate a software program's implementation is to review software development materials relevant to V&V tasks, including source code.

V. Conclusion

19. Given Cybergenetics' inference that the number of contributors to Item 2A, swabbing of pistol, is three, four, or five, this sample qualifies as a "complex mixture," complicating the interpretation of the data. Probabilistic genotyping, while intended to resolve such complications, can involve many different approaches to algorithmic design and even more approaches to implementation of those algorithms as software programs. This is demonstrated by the range of values reported between the December 5, 2019 and October 23, 2020 reports – spanning twelve orders of magnitude from 1,440 to over one quadrillion. While none of these values has been conclusively proven to be correct, or even correct as a particular implementation, access to the source code of the five open-source software programs used for the October 23, 2020 report provides opportunity for valuable insight into determining what differs between each program, why different results were calculated, and whether any of these results should be seriously questioned due to latent flaws or deficiencies. Access to TrueAllele®'s source code would be similarly helpful in assessing whether its algorithmic design and the implementation of its algorithms as a software program are appropriate for use in a criminal trial. Publicly available materials that describe the TrueAllele® software development process are not sufficient for establishing that TrueAllele® has been developed in accordance with software engineering best practices or that it has been (or even could be) verified and validated in a way that is

consistent with software engineering standards. Access to materials supporting the TrueAllele® software development and V&V processes would allow an objective evaluation as to whether public claims of the reliable operation of TrueAllele® are supported by non-public materials. Probabilistic genotyping approaches are typically used in the evaluation of the most complex/challenging test results (like those for the pistol in this case) that DNA profiling experts agree cannot be reliably evaluated with conventional statistical approaches. While it is important to consider the perspectives of experts in the areas of molecular biology, population genetics, statistics, forensic science, computer science, and software engineering it is particularly important that software engineers have the opportunity to evaluate the implementation and testing of probabilistic genotyping systems like TrueAllele®.

I declare the above is true and correct under the penalty of perjury, executed this 14th day of November, 2020, in Dayton, Ohio.

Dan E. Krane